

# ***Universal Reference RNA***

## ***A Standard for Microarray Gene Expression Experiments***

*Natalia Novoradovskaya, M.D., Ph.D.*  
*Senior Scientist*

*Metrology and Standards Needs for Gene Expression Technologies:*  
*Universal RNA Standards*

*Stanford University, March 28-29, 2003*

# Talk Agenda

## 1. Universal Reference RNA (URR):

- What is it?
- How is it used?
- Reference microarray design.

## 2. URR preparation:

- cell lines versus tissues
- selection of cell lines for genome coverage
- unique gene contribution by the individual cell lines
- RNA preparation and stability (isolation and quality control)
- validation and testing on microarrays

## 3. Microarray coverage

## 4. Comparison to another reference RNA

## 5. Lot-to-lot consistency

## What is Universal Reference RNA ?

*Universal Reference RNA (URR) is a blend of total RNA isolated from multiple cell lines representing different tissues to maximize gene expression profile*

*Universal Human Reference RNA (UHRR) - 10 human cell lines*

*Universal Mouse Reference RNA (UMRR) - 11 mouse cell lines*

*Universal Rat Reference RNA (URRR) - 14 rat cell lines*

## How Universal Reference RNA is used?

*Universal Reference RNA is used as a reference sample in any two-color hybridization experiment, where one channel (Red) is used for experimental sample and the second channel (Green) is used for a “reference”.*

*The Intensity Ratio between Red/Green channels ( $\log R/G$ ) reports the relative quantity of RNA targets in two samples.*

*Semi-quantification in combination with synthetic transcripts, which might be “spiked” into URR at fixed absolute concentrations (“SpotReport Alien” kit contains 10 different artificial synthetic sequences).*

## “Reference” Microarray Experiment Design

$$A \rightleftharpoons B \quad \frac{A}{B}$$

$$\frac{A}{R} \quad \frac{B}{R} \quad \frac{C}{R} \quad \frac{D}{R} \quad \frac{E}{R} \quad \frac{F}{R} \quad \frac{G}{R} \quad \frac{H}{R}$$

- *Reference is a common denominator between hybridizations*
- *It allows comparison of the multiple microarray experiments*

# Microarray Experiment Design

## “Applying Common Denominator”

**Macrophages:**

Normal (MM) a1AT

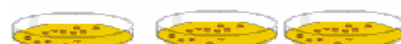
Mutant (ZZ) a1AT

Universal Human  
Reference RNA  
(Stratagene)

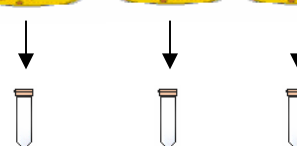
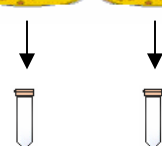
M1 M2

Z1 Z2 Z3

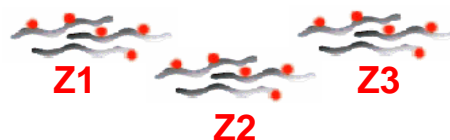
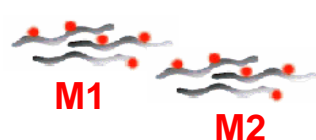
**RNA isolation:**



**Reverse  
Transcription:**



Cy5-labeled  
cDNA



**Microarray  
hybridization:**

M1  
cDNA

M2  
cDNA

Z1  
cDNA

Z2  
cDNA

Z3  
cDNA

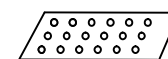
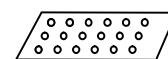
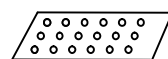
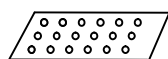
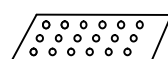
Reference  
cDNA

Reference  
cDNA

Reference  
cDNA

Reference  
cDNA

Reference  
cDNA



n=2

n=2

n=2

n=2

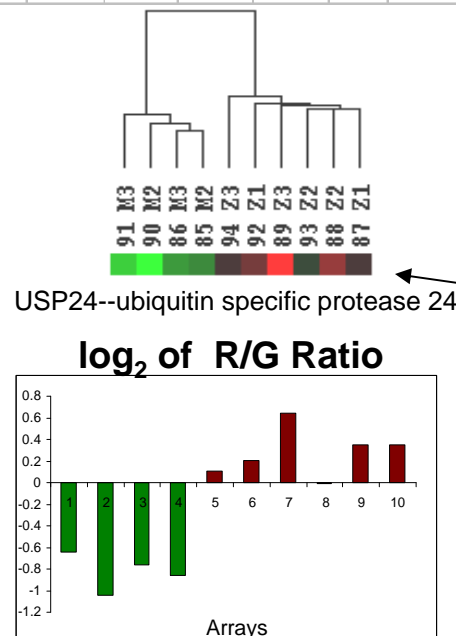
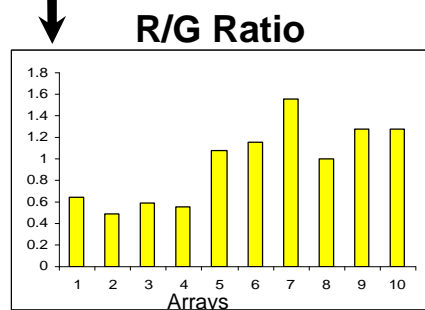
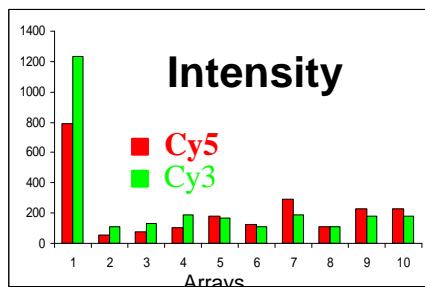
n=2

# Analysis of Microarray Data Using UHRR

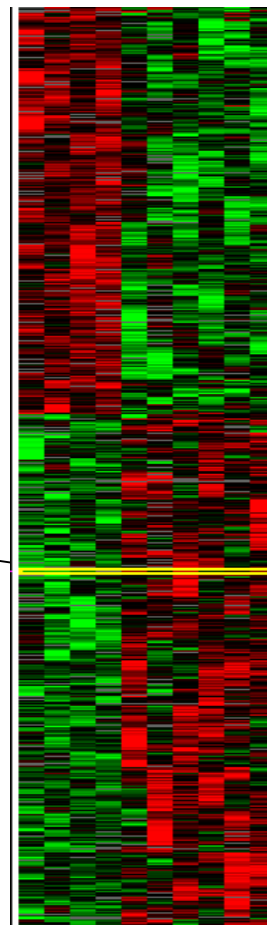
Arrays 1-10

Arrays: **1** **2** **3** **4** **5** **6** **7** **8** **9** **10**

Samples:	85 M2	90 M2	86 M3	91 M3	87 Z1	92 Z1	88 Z2	93 Z1	89 Z3	94 Z3
Intensity Cy5 (Experimental)	790	54	79	101	182	127	290	110	228	228
Intensity Cy3 (Reference-UHRR)	1235	111	134	184	169	111	186	110	178	178
R/G (Cy5/Cy3) ratio	0.64	0.49	0.59	0.55	1.08	1.15	1.56	1	1.28	1.28
log <sub>2</sub> of R/G ratio	-0.64	-1.04	-0.76	-0.9	0.11	0.2	0.64	-0	0.35	0.35



Cluster Analysis



- Fluorescence intensity depends on variations in spot size and shape, amount of spotted cDNA, labeling and hybridization conditions
- Red/Green (R/G) ratio allows comparison of relative expression levels
- Log<sub>2</sub> of R/G ratio results in symmetric distribution about zero

## Why cell lines versus tissues?

- **Availability of material**
  - *easy to grow*
- **RNA quality**
  - *RNA isolation is simple and quick, minimizing RNA degradation*
- **Lot to lot reproducibility**
  - *easy to control cell growth using standard tissue culture conditions; same passage numbers, media, temperature and harvesting method*

## Use of Pooled RNA from Different cell lines as a Reference

*Reference RNA from multiple cell lines was developed at Stanford University in Botstein / Brown Laboratory.*

*This blend of RNA from 11 human cell lines was successfully used for several years as a common reference RNA to study gene expression in multiple breast, lung and liver cancer samples.*

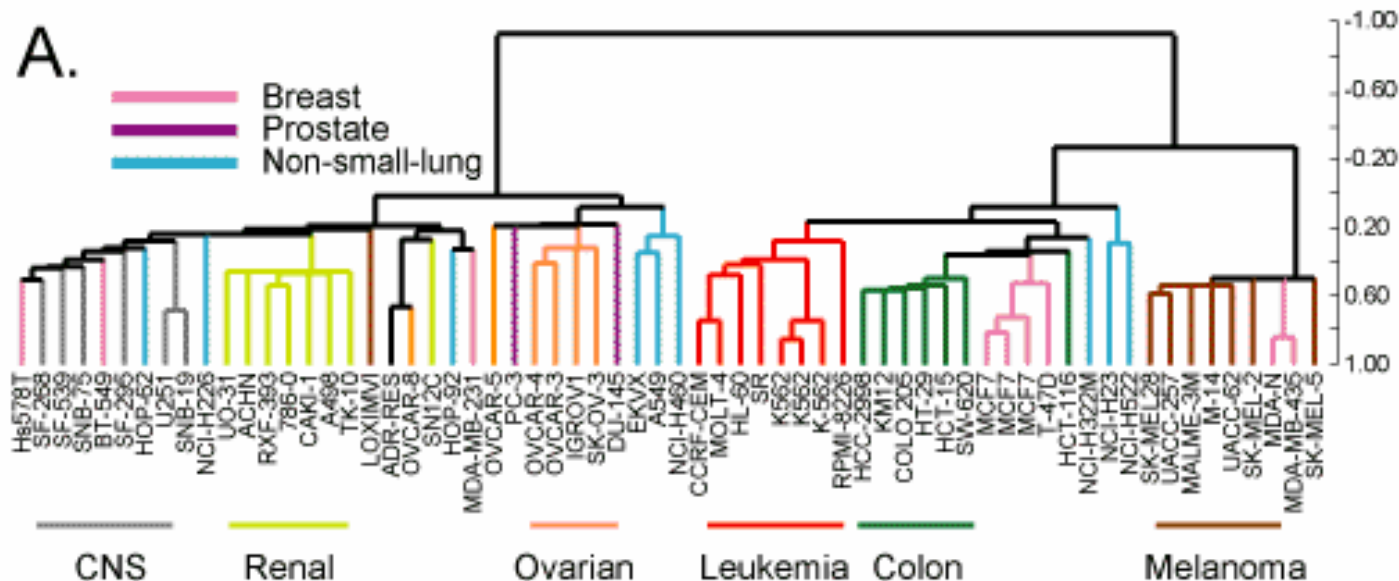
## References

- Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, Spellman P, Iyer V, Jeffrey SS, Van de Rijn M, Waltham M, Pergamenschikov A, Lee JC, Lashkari D, Shalon D, Myers TG, Weinstein JN, Botstein D, Brown PO. Systematic variation in gene expression patterns in human cancer cell lines. *Nat Genet.* 2000, Vol.3:208-9.
- Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO, Botstein D. Molecular portraits of human breast tumours. *Nature* 2000 Aug 17;406(6797):747-52.
- Whitfield ML, Sherlock G, Saldanha AJ, Murray JI, Ball CA, Alexander KE, Matese JC, Perou CM, Hurt MM, Brown PO, Botstein D. Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol Biol Cell* 2002 Jun;13(6):1977-2000.
- Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J Jr, Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Staudt LM, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000 Feb 3;403(6769):503-11.
- Chang, H.Y., J.T. Chi, S. Dudoit, C. Bondre, M. Van De Rijn, D. Botstein and P.O. Brown. 2002. Diversity, topographic differentiation, and positional memory in human fibroblasts. *Proc. Natl. Acad. Sci. USA* 99:12877-12882.
- Garber ME, Troyanskaya OG, Schluens K, Petersen S, Thaessler Z, Pacyna-Gengelbach M, van de Rijn M, Rosen GD, Perou CM, Whyte RI, Altman RB, Brown PO, Botstein D, Petersen I. Diversity of gene expression in adenocarcinoma of the lung. *Proc Natl Acad Sci U S A* 2002 Jan 22;99(2):1098.
- Chen X, Cheung ST, So S, Fan ST, Barry C, Higgins J, Lai KM, Ji J, Dudoit S, Ng IO, Van De Rijn M, Botstein D, Brown PO. Gene expression patterns in human liver cancers. *Mol Biol Cell* 2002 Jun;13(6):1929-39

## Universal Human Reference RNA (UHRR)

*Stratagene's UHRR was developed in collaboration with Botstein/ Brown's research group at Stanford University and is analogous to their common reference*

# Cell-Line Dendrogram of Gene Expression Patterns Related to the Tissue of Origin of the Cell Lines



- Cell lines derived from common tissue grouped together

Ross, D.T. et al., *Nature Genetics*, 2000, V 24, 3: 227

## Universal Human Reference RNA (UHRR)

To obtain broad gene representation, 21 human cell lines derived from different tissues were used as candidates for UHRR



Cell line selection criteria:

- gene expression pattern
- ease of growth in standard tissue culture conditions
- RNA yield and quality

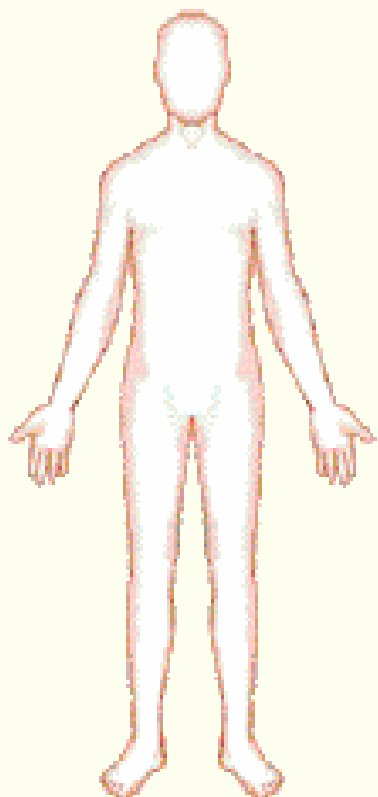


10 cell lines were selected for the final UHRR blend

# Universal Human Reference RNA Preparation

*Ten human cell lines derived from different human tissues were selected to imitate the expression profile the majority of human genes*

TISSUES → CELL LINES → RNA ISOLATION



LIVER

BRAIN

SKIN

BREAST

TESTIS

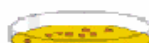
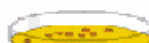
CERVIX

ADIPOCYTE

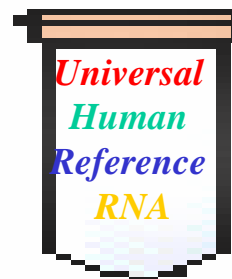
T-LYMPHOCYTE

B-LYMPHOBLAST

MACROPHAGE



*Equal quantities of total RNA from each cell line were pooled together*



- *Every gene expressed in individual cell line will be present in the Reference*

# RNA Isolation and Quality Control

## *Total RNA isolation method:*

guanidine thiocyanate/silica-based fiber matrix with DNase treatment

## *Quality Control:*

### **1. Spectrophotometry**

$OD_{260/280} > 1.8$

### **2. Formaldehyde-agarose gel electrophoresis**

18S and 28S ribosomal bands without degradation

### **3. RNase contamination testing ( 2hr incubation at 37° C)**

### **4. RNA concentration** is determined using RiboGreen Dye and by spectrophotometry

### **5. Agilent 2100 Bioanalyzer**

18S and 28S ribosomal bands without degradation,  
rRNA Ratio [28S/18S] > 1.5

### **6. Evaluation of gene representation using microarrays.**

# RNA Quality Control

Total RNA was isolated from 10 human cell lines with DNase treatment

**A**

Cell line      A 260/280

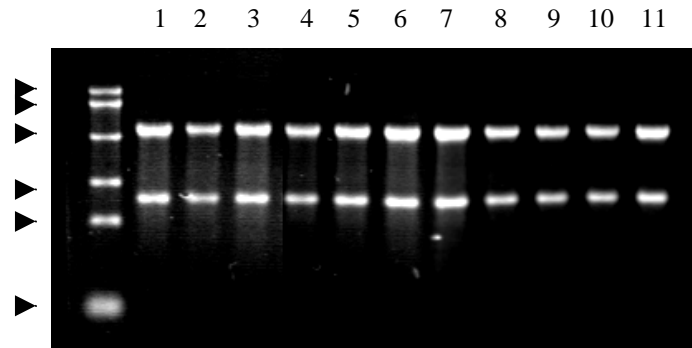
1.	1.9984
2.	1.9480
3.	1.9406
4.	1.8161
5.	1.9954
6.	1.8833
7.	1.9174
8.	1.9967
9.	1.9768
10.	1.9989

11. UHRR Pool: 1.9927

**B**

Kb

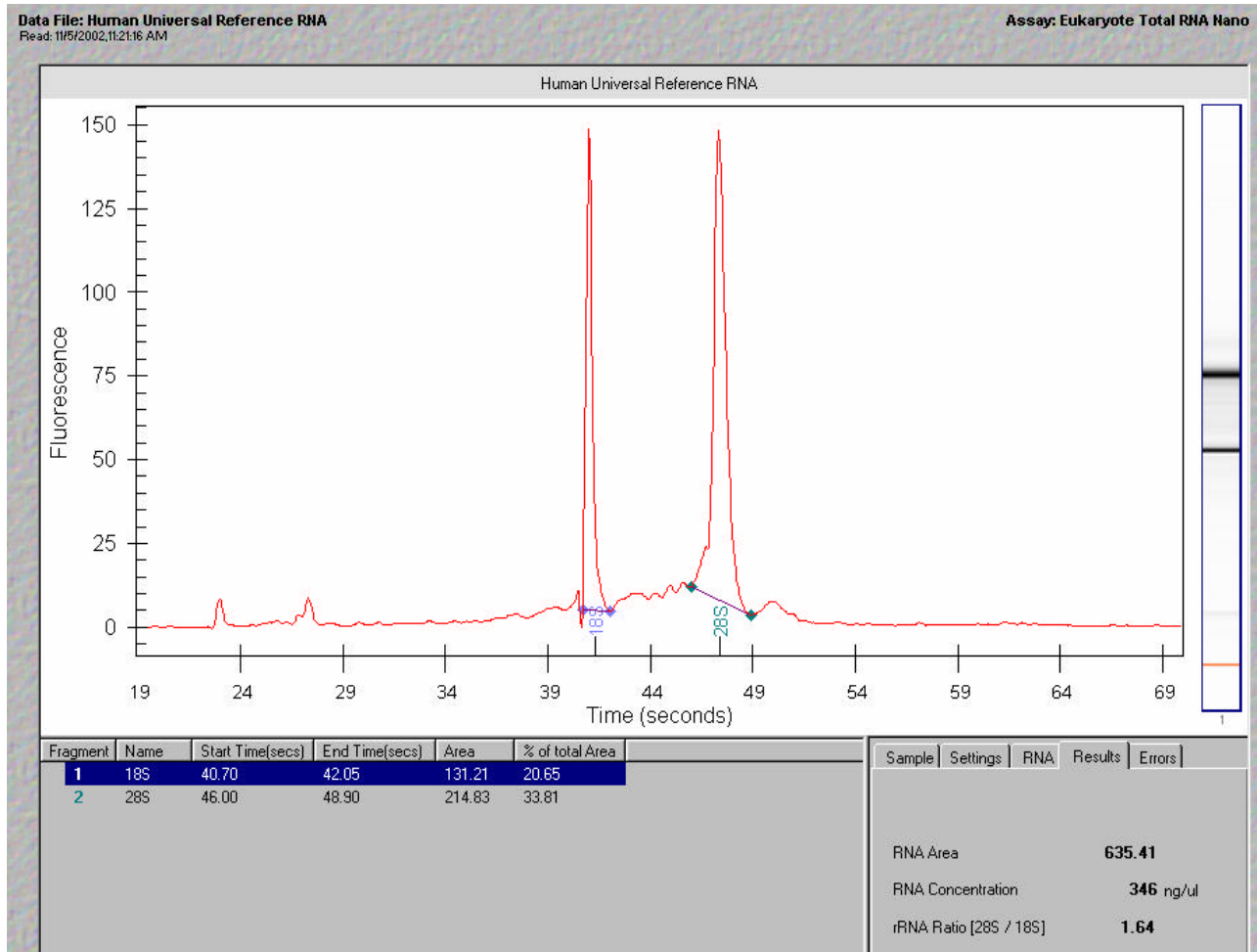
9.49  
7.46  
4.40  
2.37  
1.35  
0.24



1.25% formaldehyde-agarose gel,  
ethidium bromide staining

- $OD_{260/280} > 1.8$
- 18S and 28S ribosomal bands are without degradation

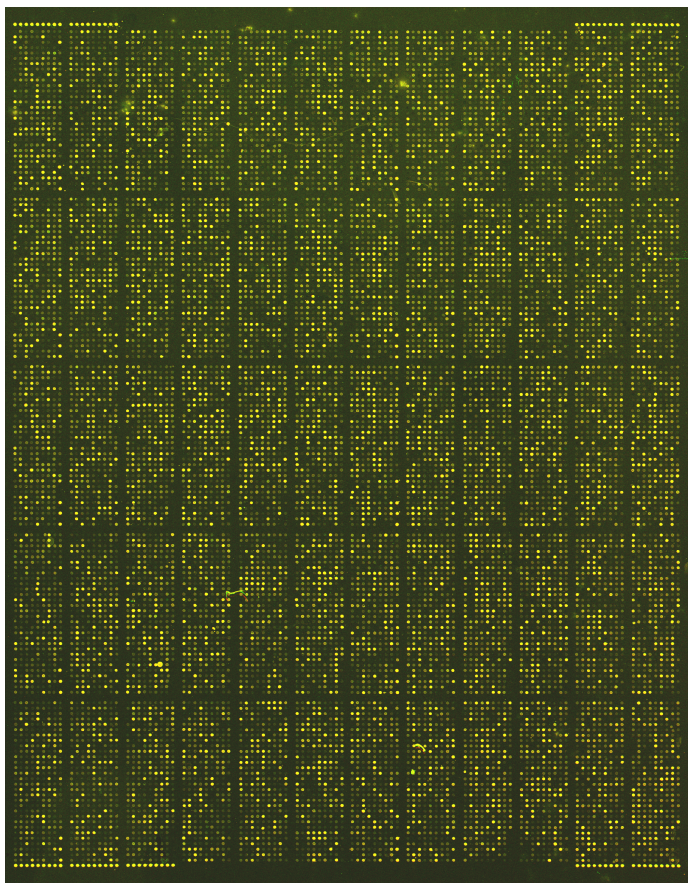
# RNA Quality Control Using Agilent Bioanalyzer



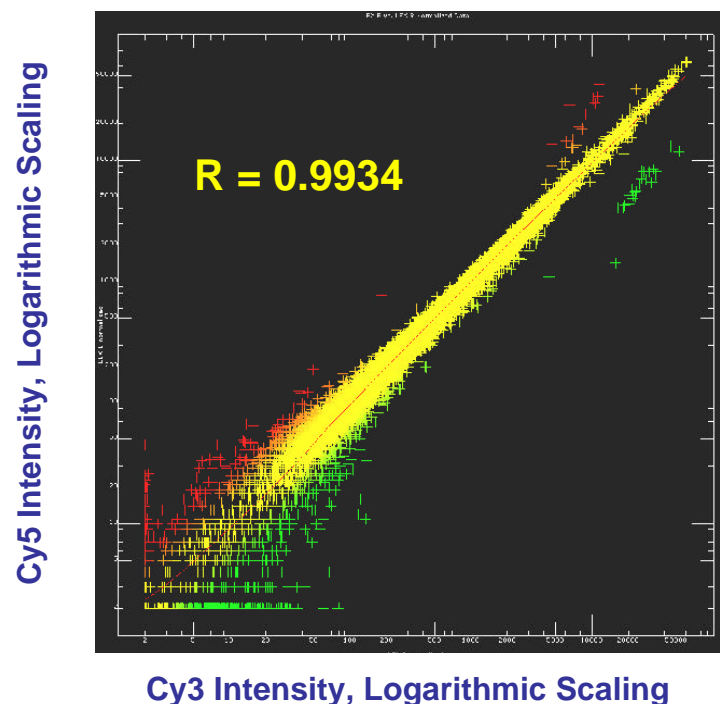
- 18S and 28S ribosomal bands without degradation
- 28S /18S ratio > 1.5

# Hybridization of UHRR to 16,000-spot Human cDNA Microarrays (Agilent)

Self-to-self hybridization  
UHRR-Cy3/UHRR-Cy5



Scatter Plot  
Normalized Data Using LOWESS/SubGrid



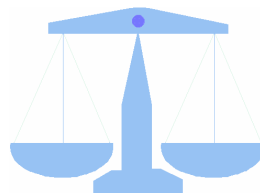
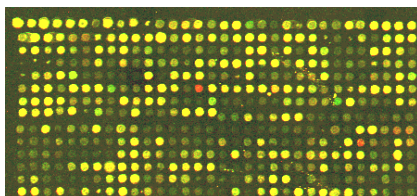
93% of spots have intensity  $> 1\times$  background  
70% of spots have intensity  $> 2\times$  background

## Microarray Coverage

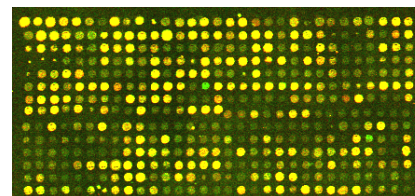
Array platform	Array number used for analysis	Spots number	Spots with Intensity -Bkgr value above:		
			zero %	1x background %	2x background %
Universal Human Reference RNA					
NCI, NIH	3	7,600	99.9	99	97
NCI, NIH	6	10,000	99.9	98	82
Agilent	4	16,000	99.9	92	68
Stanford University	6	41,000	99.9	86	63
Stanford University	11	43,000	99.9	85	60
Universal Mouse Reference RNA					
NCI, NIH	3	8,700	100	97	93
University of North Carolina	4	7,500	100	97	85
Agilent	4	8,000	100	85	65
PanVera	1	10,000	99	91	65
Universal Rat Reference RNA					
NCI, NIH	2	6,500	98	81	62
Agilent	8	14,000	99.7	86	72

## Comparison of UHRR and Stanford Common Reference RNA on 41 K Stanford cDNA Microarray

UHRR-Cy3/CRF-Cy5



CRF-Cy3/UHRR-Cy5



Total number of spots: 41,000  
Flagged spots: 5482  
Used for analysis: 35,518

Total number of spots: 41,000  
Flagged spots: 4871  
Used for analysis: 36,129

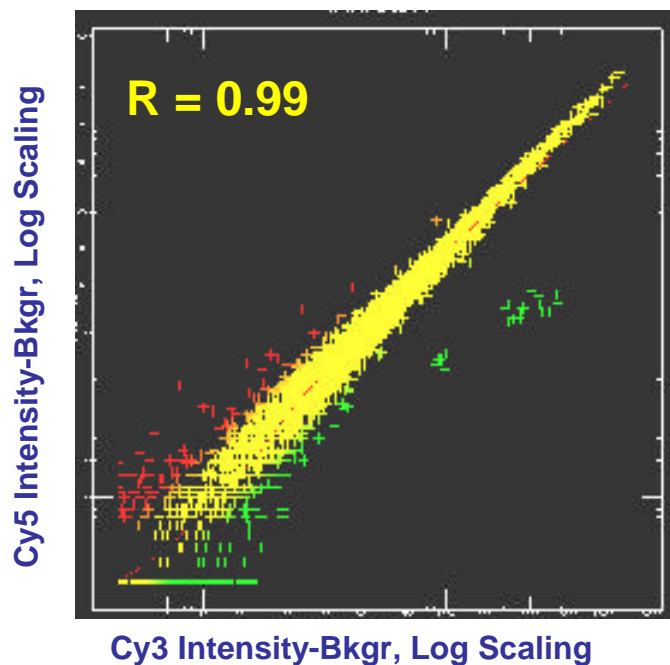
- *Stratagene Universal Human Reference RNA covers the same number of spots on 41 K microarray as Stanford common reference (~ 65 %)*
- *1649 genes expressed in Stanford reference not present in UHRR*  
*1584 genes expressed in UHRR not present in Stanford reference*

## Lot to Lot Comparison of UHRR: Hybridization of UHRR to 16,000-spots Human cDNA Microarray (Agilent)

Scatter Plots, Normalized Data (LOWESS Sub-grid)

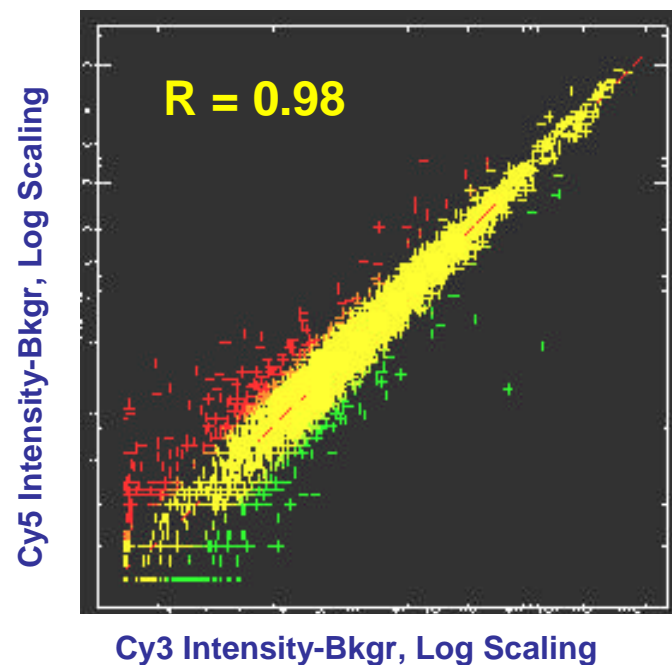
*Self -to-self hybridization*

UHRR-Cy3/ UHRR-Cy5



*Lot -to-lot comparison*

Lot#1 UHRR-Cy3/ Lot#2 UHRR-Cy5



## Organizations where URR has been adopted

- NIH
- NCI (Directors Challenge)
- TIGR
- Stanford University
- UCSF
- Boston Univ. Med Center
- Harvard Partners
- MIT
- Univ. North Carolina
- Univ. Toronto/UHN
- DKFZ
- EMBL
- Wellcome Trust – Sanger Center
- MRC
- Johnson & Johnson
- Pfizer
- Univ. Texas SWMC
- Univ. of Penn
- Sloan Kettering Memorial
- Yale Univ.
- Columbia Univ.
- Singapore Genome Institute

# Acknowledgements

## External Collaborators

### ***Stanford University:***

David Botstein  
Patrick Brown  
Michael Whitfield  
Michael Fero  
Robert Pesich  
Nicki Chin

### ***University of North Carolina:***

Charles Perou  
Jerry Usary  
Mehmet Karaca

### ***National Institute of Health:***

Jeff Green  
Mike Wilson  
Olga Aprelikova

### ***University of Texas Southwestern Medical Center***

Skip Garner  
Ryan Weil

### ***Stratagene***

Natalia Novoradovskaya  
Jeff Braman  
Scott Basehore  
Alexey Novoradovsky  
Winston Wong

### ***University of Florida:***

Mark Brantly  
Vishnu Mishra